IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BOARD OF PATENT APPEALS AND INTERFERENCES

Applicant: James Errico   Group Art Unit: 2173

Serial No.: 10/016,941   Examiner: Pillai, Namintha

Filed: December 13, 2001   Customer No.:   55648

Title: SYSTEM FOR PRESENTING AUDIO-VIDEO CONTENT

CORRECTED APPEAL BRIEF

Chernoff, Vilhauer, McClung, and Stenzel, L.L.P.
601 SW Second Avenue, Suite 1600
Portland, Oregon 97204

October 21, 2008

Mail Stop APPEAL BRIEF-PATENTS
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

BACKGROUND

This corrected brief is in response to the Notification of Non-Compliant Appeal Brief, mailed October 16, 2008, and in furtherance of the Notice of Appeal, filed in this case on June 4, 2008. The corrected brief eliminates the underline from the semi-colon at the end of limitation (c) in claim 56.

The fees required under 37. C.F.R. § 41.20(b)(2), and any required petition for extension of time for filing this brief and fees therefore, are dealt with in the accompanying TRANSMITTAL OF APPEAL BRIEF.

This brief comprises these subjects under the headings, and in the order, set forth below:

The final page of this brief bears the practitioner's signature.

REAL PARTY IN INTEREST

The real party in interest in this appeal is Sharp Laboratories of America, Inc., assignee of the captioned application.

RELATED APPEALS AND INTERFERENCES

There are no other appeals or interferences that will directly affect, be directly affected by, or have a bearing on the Board's decision in this appeal.

STATUS OF CLAIMS

A. TOTAL NUMBER OF CLAIMS IN THE APPLICATION

There are 82 claims currently pending in the application.

B. STATUS OF ALL CLAIMS

Claims canceled:	8, 35, 38, and 57
Claims withdrawn:	None
Claims pending:	1-7, 9-34, 36, 37, 39-56, and 58-86
Claims allowed:	None
Claims objected to:	None
Claims rejected:	1-7, 9-34, 36, 37, 39-56, and 58-86

C. CLAIMS ON APPEAL

Claims 1-7, 9-34, 36, 37, 39-56, and 58-86 are on appeal.

A copy of the claims on appeal is set forth in the Claims Appendix to this Brief.

STATUS OF AMENDMENTS

No amendment was filed after final rejection.

SUMMARY OF CLAIMED SUBJECT MATTER

The claimed subject matter is most broadly set forth in three independent claims.
Independent claim 1 is generally directed to a method of presenting information regarding a
video comprising a plurality of frames to a user, where the method includes six specified steps.

The first claimed step is summarizing a video, the summarization comprising a plurality of segments of the video, where each segment includes a plurality of sequential frames of that video. *See* FIG. 1, element 20; *See also* Specification at p. 4 lines 10-14; p. 6 lines 10-15. The summarization is based upon an event characterized by a semantic event that includes a sports play. *See* Specification at p. 5 line21 to p. 6 line 9. The second claimed step is displaying the summarization in a first portion of a display. *See* FIG. 1 and Specification at p. 6 lines 10-11. The third claimed step is displaying a graphical user interface on a second portion of the display, the interface sequentially indicating the relative location of each of the plurality of segments within the summarization relative to at least one other of the segments as each of the plurality of segments is displayed. *See, e.g.,* FIGS 1, 5-16; Specification at p.6 lines 13-29. Each of the plurality of segments is represented by a bounded spatial region on the second portion of said display. *Id.* The fourth claimed step is displaying to the user the relative location for a first semantic characterization of a sports play in the video using a first visual indication and displaying the relative location for a second semantic characterization of a sports play in the video using a second visual indication different from the first visual indication. *See, e.g.,* FIGS 5-9; Specification at p. 7 line 20 – p. 9 line 11. The fifth claimed step is receiving from the user, by interaction with the graphical user interface, a selection of one of said plurality of segments. *See, e.g.,* Specification at p. 8 lines 5-11. The sixth claimed step is, in response to the selection of the user, presenting a selected one of the plurality of segments and not presenting at least one other of the plurality of segments. *Id.*

Independent claim 29 is generally directed to a method of presenting information regarding a video comprising a plurality of frames to a user. The method includes six claimed

4

steps. The first step is identifying a plurality of different segments of the video, where each of the segments includes a plurality of frames of the video. *See* FIG. 1, element 20; *See also* Specification at p. 4 lines 10-14; p. 6 lines 10-15. The second claimed step is displaying, simultaneously with a segment of the video, a graphical user interface including information regarding the temporal location of one segment relative to at least one other of the segments of the video. *See, e.g.,* FIGS 1, 5-16; Specification at p.6 lines 13-29. The third claimed step is displaying in an interactive display the temporal location for a first semantic characterization of an event in the video using a first visual indication and displaying the temporal location for a second semantic characterization of an event in the video using a second visual indication that is different from the first visual indication. *See, e.g.,* FIGS 5-9; Specification at p. 7 line 20 – p. 9 line 11. The fourth claimed step is displaying to the user at least one selector by which the user may interact with the interactive display to select for viewing selective identified ones of the plurality of segments. *See* Specification at p. 7 line 26 – p. 8 line 18. The fifth claimed step is receiving user selections of identified ones of the plurality of segments. *See, e.g.,* Specification at p. 8 lines 5-11. The sixth claimed step is presenting user-selected ones of the plurality of different segments. *Id*

Independent claim 56 is generally directed to a method of presenting information regarding an audio to a user, which includes six specified steps. The first step is identifying a plurality of different segments of the audio, where each segment includes a temporal duration of audio. *See* Specification at p. 12 lines 28-33; *See also* Specification at p. 4 lines 10-14; p. 6 lines 10-15. The second step is displaying, simultaneously with the plurality of segment of audio, a graphical user interface including information regarding the temporal location of one audio

5

segment relative to at least one other audio segment. *See, e.g.,* FIGS 1, 5-16; Specification at p.6 lines 13-29. The third claimed step is displaying in an interactive display the temporal location for a first semantic characterization of an event in the audio using a first visual indication and displaying the temporal location for a second semantic characterization of an event in the audio using a second visual indication that is different from the first visual indication. *See, e.g.,* FIGS 5-9; Specification at p. 7 line 20 – p. 9 line 11.The fourth claimed step is displaying to a user at least one selector by which the user may interact with the display to select, for listening, selective identified ones of the plurality of segments. *See* Specification at p. 7 line 26 – p. 8 line 18. The fifth claimed step is receiving user selections of identified ones of the plurality of segments. *See, e.g.,* Specification at p. 8 lines 5-11. The sixth claimed step is presenting user-selected ones of the plurality of different segments. *Id.*

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

The grounds of rejection presented for review are whether claims 1-7, 9-34, 36, 37, 39-56, and 58-86 are unpatentable under 35 U.S.C. §103(a) over the combination of Christel et al., "Adjustable Filmstrips and Skims as Abstractions for a Digital Video Library" (hereinafter Christel) in view of Vasconcelos et al., "Bayesian Modeling of Video Editing and Structure: Semantic Features for Video Summarization and Browsing" (hereinafter Vasconcelos) and in further view of Ahmad et al., U.S. Patent No. 6,880,171 (hereinafter Ahmad).

ARGUMENT

1.   Rejection of claims 1-7, 9-34, 36, 37, 39-56, and 58-86.

The Examiner rejected claims 1-7, 9-34, 36, 37, 39-56, and 58-86 as being unpatentable over the combination of Christel in view of Vasconcelos, and in further view of Ahmad. Independent claim 1 can be best understood by referencing Figure 5 of the applicant's disclosure. A user is presented with a graphical user interface (GUI) next to a display that presents a customizable video summary to a user. The GUI presents a timeline 30 that shows the relative locations of *different types* of interesting segments included in a summary presented in the display. For example, as seen in FIG. 5, segments having slam dunk plays may be visually annotated in a different manner than are segments that have three point shots, etc. A user viewing the summary, who wishes to view three point shots, may use a selector such as a scroll bar 56 to select a three point segment, following which the video presentation in the display will move to the location requested by the user. The features described above, that distinguish claim 1 over the prior art, are contained in the following specified limitations:

(1)   displaying said summarization in a first portion of a display;

(2)   displaying a graphical user interface on a second portion of said display, said interface sequentially indicating the relative location of each of said plurality of segments within said summarization relative to at least one other of said segments as each of said plurality of segments is displayed, each of said plurality of segments represented by a bounded spatial region on said second portion of said display;

(3)   displaying to a user said relative location for a first semantic characterization of a said sports play in said video using a first visual indication and displaying said relative location for a second semantic characterization of a said sports play in said video using a second visual indication different from said first visual indication;

(4)   receiving from said user, by interaction with said graphical user interface, a selection of one of said plurality of segments; and

(5)   in response to said selection, presenting a selected one of said plurality of segments and not presenting at least one other of said plurality of segments.

With these limitations delineated, the applicant will address the Examiner's rejection of claim 1 in view of the combination of Christel, Vasconcelos, and Ahmad.

Christel, the primary reference, discloses a system for presenting video skims in which a user may enter a specific query to which certain frames of a video are "matched." The video skim is constructed by (1) using a query from a user to identify matching key frames in a selected video, and (2) based upon those matching frames, constructing a summarization that builds video segments around each of the matching frames. The first of these steps is accomplished by matching words in the query to words in descriptive text of the video, constructed from either a speech recognition module or annotations to the video, such as close captioning, or both. *See, e.g.,* Christel at p. 3 col. 2. Because the query-matching module of

Christel relies so heavily upon a textual description of the video, Christel discloses that the system is limited to an Informedia video collection that includes news and documentaries, i.e. genres of video for which the textual descriptions not only actually describe the content of what is visually presented in the video, but is timed to coincide with the segments they describe. *See, e.g.,* Christel at p. 1 col. 1 section 1, par. 1 (describing the system being applied to news and documentary videos); *See also Id.* at p. 3 col. 2, section 3 par. 3 (stating that the retrieval engine relies upon matching words in a query to descriptive text "timed tightly" to the video segment). Christel also discloses a selector by which a user may adjust the compression ratio used in constructing the summary, i.e., the compression ratio determines the size of the segments that are built around the matching frame locations.

(a)   The prior art fails to disclose the steps of a user selecting one of the segments and presenting that segment without presenting at least one other segment.

Christel discloses a user interface for a video summary that presents, in two separate bars or lines, locations of individual frames matching a query, and segments generated around those matching frames the selected compression ratio. *See* FIGS 5 and 6 of Christel. Thus, as correctly noted by the Examiner, this latter bar shows a "plurality of segments [each] represented by a bounded spatial region on said display." *See* Christel at FIGS 5 and 6. The Examiner, however, leaps to the conclusion that Christel's indication that a user can adjust the compression ratio, thereby creating *new, different-sized bounded segments* each centered around a respective match location, discloses the limitations of a user "selecting" one of the plurality of segments and subsequently being presented with the selected segment and not another segment. Claim 1, however, does not permit such an interpretation, because the Examiner ignores the antecedent

relationship between the selected and presented "said" segment to the *previously defined* segments that are "represented by a bounded spatial region on said display." Christel's disclosure of a compression-ratio selector, by which a user may begin anew to reconstruct another set of bounded segments to be included in a summary, is not a selection of any of the bounded segments defined by the claims. If anything, a user adopting a new compression ratio *is a rejection of those segments*. Once the new compression ratio is input, those previous segments are *replaced* by new segments, *all of which* are then shown in the summary. Thus, Christel does not disclose a user selecting one of the summary segments, and then having that segment presented to the user to the exclusion of at least one other segment.

The Examiner's mistake lies in not attaching significance to the claimed antecedent reference in the limitations of "receiving from said user, by interaction with said graphical user interface, a selection of one of *said* plurality of segments" and "in response to said selection, presenting a selected one of *said* plurality of segments." Once the Examiner has read the initial limitation of a "plurality of segments" each "represented by a bounded spatial region" on a display, upon Christel's segments shown in FIG. 5, then the Examiner must show a disclosure in Christel that one of *those* segments, with *those* bounded spatial regions, is selected and presented to the exclusion of other displayed bounded segments. The Examiner has not done so. Instead, the Examiner cites to an irrelevant portion of Christel that allows a user to create *new* segments, with *new* bounded regions, and present *all* of the new segments.

The Examiner's rejection of claim 1 is premised on the assumption that Christel discloses the limitations of "receiving from said user, by interaction with said graphical user interface, a selection of one of said plurality of segments" and "in response to said selection, presenting a selected one of said plurality of segments and not presenting at least one other of said plurality of

10

segments." Because Christel fails to disclose this limitation, the Examiner's rejection is improper.

(b)   The prior art fails to disclose the limitation of "displaying to a user said relative location for a first semantic characterization of a said sports play in said video using a first visual indication and displaying said relative location for a second semantic characterization of a said sports play in said video using a second visual indication different from said first visual indication."

The Examiner concedes that this limitation is not disclosed by Christel, but instead argues that it is an obvious modification of Christel in view of Vasconcelos and Ahmad. The applicant respectfully argues that the Examiner is mistaken.

At the outset, to the extent that Christel does disclose "summarizing a video, said summarization comprising a plurality of segments of said video based upon an event", the "event" that forms the basis for selecting the plurality of segments to include in the summary is unknown to all but the contemporaneous user, hence is both unpredictable, and *already tailored to the user's specific query*. In fact, Christel touts these features:

> Just as we modified filmstrips so that match locations were taken into account, so video skims were adjusted from early work to emphasize the audio and video surrounding match locations. Rather than being pre-computed, these new style video skims are generated dynamically so that context can be used to assemble better skims, e.g., following a query the skim will be assembled to emphasize the locations in the video where match locations are found.

See Christel at p. 4 col. 2 lines 38-47.

This primary reference notes, however, that merely constructing a summary of a video by extracting the segments in a video that best match a user's contemporaneous interests is insufficient to hold the user's interest in the skim; mere extraction of these segments tended to produce an aesthetically displeasing, choppy and unsynchronized video. *See* Christel at p. 4 col. 2 lines 16-31. To improve the fluidity of the summary presentation, Christel proposed a method of constructing a skim by expanding segments around match locations, where the length of each segment was determined by a combination of (1) user input as to the compression ratio for the summary as a whole; and (2) "goodness values" calculated for automatically-generated segment cutpoints. *See Id.* at p. 4 col. 2 line 45 to p. 5 col. 1 line 4. Specifically, given user input of a query and a desired compression ratio,

> [t]he skim is initialized to consist of sequences containing any of the given match locations, merging sequences which occur very close together. The sequences in the skim are then expanded: the sequence endpoint with the worst goodness rating is extended out to the next cutpoint, thus embedding that bad cutpoint into the skim. This process repeats until the target skim size is reached.

See Id. at p. 5 col. 1 lines 14-21.

Given the inherent trade-off between the user-selected compression ratio and the fluidity of the skim presentation, Christel shows a user interface that roughly communicates to a user the marginal benefit of decreasing the selected compression ratio (increasing the length of the summary), and allows a user to adjust the compression ratio accordingly. In FIG. 5, for example, Christel shows a user interface that, in addition to playing the desired skim, shows two bars. The first bar indicates the relative location of matching frames in the video being skimmed, while the

second bar indicates the relative location of the segments automatically constructed around those

matching frames. From these two bars, a user can estimate the marginal improvement in

presentation by decreasing the selected compression ratio. For example, in FIG. 5 a majority of

the match locations to the exemplary query are found near the beginning of the video, and many

of the segments are interrupted by only a short interval. Thus, it would be reasonable to assume

that marginally decreasing the compression ratio would achieve a proportionally greater benefit

in presentation fluidity. This is confirmed by FIG. 6, where, by decreasing the amount of

compression from 20% (5:1 compression) to 40% (5:2 compression), the number of breaks

between segments was reduced from 12 to 5. Moreover, FIG. 6 shows a slider allowing the

viewer to incrementally adjust the compression ratio using feedback from the segment and match

point location bars.

Therefore, although Christel discloses the claimed step of "displaying a graphical user

interface on a second portion of said display, said interface sequentially indicating the relative

location of each of said plurality of segments within said summarization relative to at least one

other of said segments as each of said plurality of segments is displayed" as recited in

independent claim 1, Christel does so *solely for the purpose of providing statistical feedback to*

the user as to the marginal benefits received in exchange for the cost of further decreasing the

compression ratio, increasing the length of the summary. The graphical user interface of Christel

is not intended to distinguish the relative temporal locations of *different types of* semantic

content, nor would it be used for such a purpose because the summary of Christel is *already*

constructed in response to a specific user inquiry as to the type of semantic content the user

wishes to see.

Nonetheless, the Examiner asserts that Vasconcelos and Ahmad indicate the obviousness of including this type of redundancy. Neither reference, however, either teaches this limitation or provides a reason for modifying Christel to provide it.

Vasconcelos teaches a method of automatically identifying a high-level semantic domain of a film, i.e. whether the film is of a particular genre, e.g. action, romance, etc. Drawing on the observation that different genres of movies have different image characteristics, e.g. that dramas tend to be heavy on dialogue and close-ups of actors' faces, while action films tend to employ fast cuts with fewer close-ups of actors, Vasconcelos describes that a video may be characterized by four timelines, shown in FIG. 2, that each show the temporal locations of respective ones of four characteristic *statistical* types of video *shots*, i.e. close-up shots, fast-cut shots, crowd shots, and natural settings. *See* Vasconcelos, FIG. 2; *see also Id.* at p. 153, section 1, par. 2 (stating that the method identifies the structure of shots in a video from which semantic attributes can be inferred). Vasconcelos discloses that by comparing these timelines for a video, a viewer can infer the semantic type of film being shown. Stated differently, Vasconcelos discloses the use of timelines, not for distinguishing among a plurality of types of semantic events in a video, but instead for distinguishing among a plurality of types of visual composition of shots. *See* Vasconcelos at p. 154 sec. 2 (describing a Bayesian inference process where detected *structural features* of movies are mapped to a graph, from which inferences can be made as to semantic content). The genre of film can then be inferred from viewing the timelines together, e.g. if the timelines show more close-up shots than fast-cut scenes (both characteristics of the shot composition, as opposed to a type of semantic event in the video itself), then a viewer can infer a likelihood that the video is of the drama genre than the action genre.

14

With this in mind, the Examiner's assertion that Vasconcelos discloses identifying specific semantic events and displaying identified data through visual indications in a timeline is facially incorrect. Not only do the timelines of Vasconcelos fail to display specific semantic events in a video, but Vasconcelos seems to disclaim that the system disclosed therein is capable of doing so. *See* Vasconcelos at p. 153 section 1 paragraph 4 (describing the practical outcome of the theoretical approach outlined in the paper as being "generic", i.e. only characterizing the *domain* of movies); *See also Id.* at p. 156, section 5.3 ("The characterization is not fine enough to [automatically] distinguish between The River Wild and Ghost and the Darkness"). Therefore neither Vasconcelos nor the primary reference, Christel, discloses the limitation of "displaying said relative location for a first *semantic characterization of a said play* in said video using a first visual indication and displaying said relative location for a second *semantic characterization of a said play* in said video using a second visual indication different from said first visual indication."

Nor does the tertiary reference, Ahmad, disclose this limitation. Ahmad discloses a browser for audiovisual content where a user can view summary information related to available content. In a specific embodiment, noted by the Examiner, Ahmad discloses a window showing, as an example, "news programs" available for viewing where any currently viewed news program is shaded in one color while news programs that have already been viewed are shaded in another color. *See* Ahmad at col. 16 lines 54-65. Presumably, were the window showing "action movies" or "documentaries" the window could be similarly marked to shade, for example, any currently viewed documentary one color and previously viewed documentaries another color. Thus, the different colored shadings, as taught by Ahmad, are not indicative of any semantic content *in* the video; rather, the differing visual indications are merely indicative of the

statistical property of whether that viewer is either currently watching the program (shading in one color), has previously watched the program (shading in another color), or neither (no shading). The applicant further notes that the post-facto marking of content as being either watched or not watched cannot indicate anything meaningful about the events in a video created long before the user had the opportunity to watch the program.

The term "semantic event" relates to the *meaning* of an event, and more specifically, the claim limitation of a "semantic characterization of a play" (or event) in a video relates to a meaning of a particular play or event portrayed. For example, if the video is of a basketball game, a type of semantic characterization of a play (event) in the video might include slam dunks, fast breaks, fouls, and injuries. If the video is an action movie, types of semantic characterizations of events in the video might include car chases, explosions, and gunfights. Even a cursory reading of Ahmad shows that it fails to disclose the limitation of "displaying said relative location for a first semantic characterization of a play (or event) in said video using a first visual indication and displaying said relative location for second semantic characterization of a play (or event) in said video using a second visual indication different from said first visual indication."

In view of the respective disclosures of Christel, Vasconcelos, and Ahmad, each previously described, the Examiner's rejection of independent claim 1 as being obvious over the combination of these references is deficient because none of the cited references disclose using visual indicia in a graphical interface to indicate respective types of *semantic content* depicted in a video being summarized. Instead, each reference uses visual indicia to show statistical or structural properties of either the video (e.g. Vasconcelos' timelines showing types of structure of shots; Ahmad's colors indicating the statistical feature of whether a video has been watched)

or a summary of a video (e.g. Christel's video scroll bars showing match locations and segment locations used in a video summary, relative to the summarized video).

Moreover, one of ordinary skill in the art would not make the combination suggested. If, for example, the Examiner is arguing for a substitution of Vasconcelos' timelines (as modified to include Ahmad's different colors for different shot-types) for Christel's scroll bars, then such a substitution would frustrate the very purpose of Christel's user interface, which is to provide the user feedback as to the marginal benefit of making the summary a little longer. On the other hand, if the Examiner is suggesting that Christel's user interface be modified to include, in addition to the scroll bars, the timelines of Vasconcelos, the Examiner fails to provide a motive for doing so; a user of Christel's system already knows the genre of the video being summarized, and is being presented with a summary specifically constructed in response to a query as to the type of content desired to be seen. A user of Christel's summary has no need for the timelines of Vasconcelos because there is no need to infer, using a Bayesian model or otherwise, what content is being presented when the content presented already matches a specific query.

The Examiner's conclusion that the limitation in claim 1, that it would be obvious, in light of Vasconcelos and Ahmad, to modify Christel to arrive at the limitation of "displaying to a user said relative location for a first semantic characterization of a said sports play in said video using a first visual indication and displaying said relative location for a second semantic characterization of a said sports play in said video using a second visual indication different from said first visual indication" lacks support in the prior art.

Dependent claims 2-7 and 9-28 depend from independent claim 1 and are therefore also distinguished over the cited prior art. Independent claims 29 and 56 include the limitations of "displaying in an interactive display said temporal location for a first semantic

17

characterization of an event in said video using a first visual indication and displaying said temporal location for a second semantic characterization of an event in said video using a second visual indication different from said first visual indication", "displaying to a user at least one selector by which said user may interact with said interactive display to select for viewing selective identified ones of said plurality of segments", "receiving user-selections of identified ones of said plurality of segments" and "presenting user-selected ones of said plurality of different segments." Therefore claims 27 and 56, as well as their respective dependent claims 28-34, 36, 37, 39-55, and 57-86 also distinguish over the cited prior art.

2. Rejection of claims 6, 7, and 9-11.

These dependent claims are directed to a disclosed feature of the applicant's described interface where a user may select for presentation one of the segments included in the summary by selecting a point within the "bounded spatial region" on the interface corresponding to the segment. (claim 6). Specifically, the specification enables multiple different responses to this selection. First, if a user selects a point in a bounded spatial region, the presentation of the segment corresponding to that bounded spatial region may "snap to" the first frame of the segment. (claims 7 and 11). Second, if a user selects a point in the bounded spatial region, presentation of the segment may begin mid-segment at a frame corresponding to the location the user selected within the bounded spatial region. (claim 9). Moreover, the GUI may include a selector by which the user may select which of these modes will be used. (claim 10).

The Examiner alleges that each of these limitations are disclosed in FIGS 5 and 6 of Christel. The applicant has examined these figures, along with the text accompanying these figures, and cannot read in Christel any disclosure of the features claimed in these dependent claims.

18

CONCLUSION

The Examiner's respective rejections of claims 1-7, 9-34, 36, 37, 39-56, and 58-86 should

be reversed, and the claims should be found patentable.

Respectfully submitted,

Kurt Rohlfs
Reg. No. 54,405
Attorney for Applicant
Telephone: (503) 227-5631

CLAIMS APPENDIX

1.   A method of presenting information regarding a video comprising a plurality of frames comprising:

(a)   summarizing a video, said summarization comprising a plurality of segments of said video, based upon an event characterized by a semantic event that includes a sports play, where each of said segments includes a plurality of sequential frames of said video;

(b)   displaying said summarization in a first portion of a display; and

(c)   displaying a graphical user interface on a second portion of said display, said interface sequentially indicating the relative location of each of said plurality of segments within said summarization relative to at least one other of said segments as each of said plurality of segments is displayed, each of said plurality of segments represented by a bounded spatial region on said second portion of said display;

(d)   displaying to a user said relative location for a first semantic characterization of a said sports play in said video using a first visual indication and displaying said relative location for a second semantic characterization of a said sports play in said video using a second visual indication different from said first visual indication; and

(e)   receiving from said user, by interaction with said graphical user interface, a selection of one of said plurality of segments; and

(f)   in response to said selection, presenting a selected one of said plurality of segments and not presenting at least one other of said plurality of segments.

2.   The method of claim 1 wherein said first and second semantic characterizations of a said sports play temporally overlap in said summarization.

3.   The method of claim 1 wherein said graphical user interface includes a generally rectangular region where each of said plurality of segments is indicated within said generally rectangular region.

4.   The method of claim 1 wherein the size of each of said plurality of segments is indicated in a manner such that said plurality of segments with a greater number of frames are larger than said plurality of segments with a lesser number of frames.

5.   The method of claim 4 wherein the size of the regions between each of said plurality of segments is indicated in a manner such that said regions between with a greater number of frames are larger than said plurality of segments with a lesser number of frames.

6.   The method of claim 4 where said user selects one of said plurality of segments by interacting with said graphical user interface at a point within the displayed bounded spatial region associated with the selected one of said plurality of segments.

7.   The method of claim 6 wherein presentation of a selected one of said plurality of segments begins at the first frame of said segment irrespective of which point within said displayed bounded spatial region that said user interacted with.

8 (Canceled).

9.   The method of claim 6 wherein presentation of a selected one of said plurality of segments begins at a frame of said segment temporally corresponding to the point within said displayed bounded spatial region that said user interacted with.

10.   The method of claim 6 including a selector by which said user may alternatively select a chosen one of (i) presentation of a selected one of said plurality of segments beginning at the first frame of said segment irrespective of which point within said displayed bounded spatial region that said user interacted with; and (ii) presentation of a selected one of said plurality of segments beginning at a frame of said segment temporally corresponding to the point within said displayed bounded spatial region that said user interacted with.

11.   The method of claim 7 including a user-moveable scroll bar on said graphical user interface indicating the relative temporal location of currently-presented frames of said summary, wherein said user selects one of said plurality of segments by moving said scroll bar over the selected one of said plurality of segments, and where said scroll bar snaps to the beginning of the selected one of said plurality of segments.

12.   The method of claim 1 wherein at least two of said plurality of segments are temporally overlapping.

13.   The method of claim 12 wherein said temporally overlapping segments are visually indicated in a manner such that each of said overlapping segments are independently identifiable.

14.   The method of claim 1 wherein a user selects a portion of said video not included within said plurality of segments, wherein in response thereto, said system presents one of said plurality of segments.

15.   The method of claim 14 wherein said one of said plurality of segments is the segment most temporally adjacent to said portion of said video.

16.   The method of claim 14 wherein said one of said plurality of segments is the next temporally related segment.

17.   The method of claim 14 wherein said one of said plurality of segments is the previous temporally related segment.

18.   The method of claim 1 wherein a user selects a portion of said video included within said plurality of segments, wherein in response thereto, said system presents said portion of said video from the start thereof.

19.   The method of claim 1 wherein a user selects a portion of said video not included within said plurality of segments, wherein in response thereto, said system presents one of said plurality of segments, and wherein said user selects a portion of said video included within said plurality of segments, wherein in response thereto, said system presents said portion of said video within said plurality of segments.

20.   The method of claim 1 wherein a user selects a portion of said video not included within said plurality of segments, wherein in response thereto, said system presents one of said plurality of segments, and wherein said user selects a portion of said video included within said plurality of segments, wherein in response thereto, said system presents said portion of said video within said plurality of segments starting from the beginning thereof.

21.   The method of claim 1 wherein a user selects a portion of said video not included within said plurality of segments, wherein in response thereto, said system presents said selected portion not included within said plurality of segments, and wherein after presenting said selected portion not included within said plurality of segments presents said selected plurality of segments in temporal order without portions of said video not included within said plurality of segments, and wherein said user selects a portion of said video included within said plurality of segments, wherein in response thereto, said system presents said portion of said video within said plurality of segments.

22.   The method of claim 1 wherein said temporal information is hierarchical and is displayed in such a manner to retain a portion of its hierarchical structure.

23.   The method of claim 1 wherein said temporal information relates to overlapping time periods and said temporal information is displayed in such a manner to maintain the differentiation of said overlapping time periods.

24.   The method of claim 1 wherein said temporal information is displayed within a time line, wherein the temporal information is presented in a plurality of layers in a direction perpendicular to the length of said time line.

25.   The method of claim 1 wherein said temporal information is displayed within a time line, wherein textual information is included within said time line.

26.   The method of claim 1 wherein said temporal information is displayed within a time line, wherein additional textual information is displayed upon selecting a portion of said time line.

27.   The method of claim 1 wherein said temporal information is displayed together with a time line, wherein additional textual information is displayed together with selecting a portion of said time line.

28.   The method of claim 1 wherein said temporal information is displayed within a time line, wherein additional audio annotation is presented upon presenting a portion of said time line.

29.   A method of presenting information regarding a video comprising a plurality of frames comprising:

(a)   identifying a plurality of different segments of said video, where each of said segments includes a plurality of frames of said video;

(b)   displaying, simultaneously with a said segment of said video, a graphical user interface including information regarding the temporal location of a said segment relative to at least one other of said segments of said video;

(c)   displaying in an interactive display said temporal location for a first semantic characterization of an event in said video using a first visual indication and displaying said temporal location for a second semantic characterization of an event in said video using a second visual indication different from said first visual indication;

(d)   displaying to a user at least one selector by which said user may interact with said interactive display to select for viewing selective identified ones of said plurality of segments;

(e)   receiving user-selections of identified ones of said plurality of segments; and

(f)   presenting user-selected ones of said plurality of different segments.

30.   The method of claim 29 wherein said graphical user interface includes a generally rectangular region where each of said plurality of segments is indicated within said generally rectangular region.

31.   The method of claim 29 wherein the size of each of said plurality of segments is indicated in a manner such that said plurality of segments with a greater number of frames are larger than said plurality of segments with a lesser number of frames.

32.   The method of claim 31 wherein the size of the regions between each of said plurality of segments is indicated in a manner such that said regions between with a greater number of frames are larger than said plurality of segments with a lesser number of frames.

33.   The method of claim 29 further comprising an indicator that indicates the current position within said temporal information of a currently displayed portion of said video.

34.   The method of claim 33 wherein said indicator changes location relative to said temporal information as the portion of said currently displayed portion of said video changes.

35 (Canceled).

36.   The method of claim 29 further comprising

(a)   indicating with an indicator the current position within said temporal information of a currently displayed portion of said video; and

(b)   modifying the position of said indicator within said temporal information which modifies the displayed portion of said video.

37.   The method of claim 36 wherein said indicator is modified to a portion of said video that is not included within said plurality of segments.

38 (Canceled).

39.   The method of claim 29 wherein at least two of said plurality of segments are temporally overlapping.

40.   The method of claim 39 wherein said temporally overlapping segments are visually indicated in a manner such that each of said overlapping segments are independently identifiable.

41.   The method of claim 29 wherein a user selects a portion of said video not included within said plurality of segments, wherein in response thereto, said system presents one of said plurality of segments.

42.   The method of claim 41 wherein said one of said plurality of segments is the segment most temporally adjacent to said portion of said video.

43.   The method of claim 41 wherein said one of said plurality of segments is the next temporally related segment.

44.   The method of claim 41 wherein said one of said plurality of segments is the previous temporally related segment.

45.   The method of claim 29 wherein a user selects a portion of said video included within said plurality of segments, wherein in response thereto, said system presents said portion of said video from the start thereof.

46.   The method of claim 29 wherein a user selects a portion of said video not included within said plurality of segments, wherein in response thereto, said system presents one

of said plurality of segments, and wherein said user selects a portion of said video included within said plurality of segments, wherein in response thereto, said system presents said portion of said video within said plurality of segments.

47. The method of claim 29 wherein a user selects a portion of said video not included within said plurality of segments, wherein in response thereto, said system presents one of said plurality of segments, and wherein said user selects a portion of said video included within said plurality of segments, wherein in response thereto, said system presents said portion of said video within said plurality of segments starting from the beginning thereof.

48. The method of claim 29 wherein a user selects a portion of said video not included within said plurality of segments, wherein in response thereto, said system presents said selected portion not included within said plurality of segments, and wherein after presenting said selected portion not included within said plurality of segments presents said selected plurality of segments in temporal order without portions of said video not included within said plurality of segments, and wherein said user selects a portion of said video included within said plurality of segments, wherein in response thereto, said system presents said portion of said video within said plurality of segments.

49. The method of claim 29 wherein said temporal information is hierarchical and is displayed in such a manner to retain a portion of its hierarchical structure.

50. The method of claim 29 wherein said temporal information relates to overlapping time periods and said temporal information is displayed in such a manner to maintain the differentiation of said overlapping time periods.

51. The method of claim 29 wherein said temporal information is displayed within a time line, wherein the temporal information is presented in a plurality of layers in a direction perpendicular to the length of said time line.

52. The method of claim 29 wherein said temporal information is displayed within a time line, wherein textual information is included within said time line.

53. The method of claim 29 wherein said temporal information is displayed within a time line, wherein additional textual information is displayed upon selecting a portion of said time line.

54. The method of claim 29 wherein said temporal information is displayed together with a time line, wherein additional textual information is displayed together with selecting a portion of said time line.

55. The method of claim 29 wherein said temporal information is displayed within a time line, wherein additional audio annotation is presented upon presenting a portion of said time line.

56.   A method of presenting information regarding audio comprising:

(a)   identifying a plurality of different segments of said audio, where each of said segments includes a temporal duration of said audio;

(b)   displaying simultaneously with said segment of said audio a graphical user interface including information regarding the temporal location of a said segment relative to at least one other of said segment of said audio;

(c)   displaying in an interactive display said temporal location for a first semantic characterization of an event in said audio using a first visual indication and displaying said temporal location for a second semantic characterization of an event in said audio using a second visual indication different from said first visual indication;

(d)   displaying to a user at least one selector by which said user may interact with said display to select for listening selective identified ones of said plurality of segments;

(e)   receiving user-selections of identified ones of said plurality of segments; and

(f)   presenting user-selected ones of said plurality of different segments.

57 (Canceled).

58.   The method of claim 56 further comprising

(a)   indicating with an indicator the current position within said temporal information of a currently displayed portion of said audio; and

(b)   modifying the position of said indicator within said temporal information which modifies the displayed portion of said audio.

31

59.   The method of claim 58 wherein said indicator is modified to a portion of said audio that is not included within said plurality of segments.

60.   The method of claim 56 wherein at least two of said plurality of segments are temporally overlapping.

61.   The method of claim 60 wherein said temporally overlapping segments are visually indicated in a manner such that each of said overlapping segments are independently identifiable.

62.   The method of claim 56 wherein a user selects a portion of said audio not included within said plurality of segments, wherein in response thereto, said system presents one of said plurality of segments.

63.   The method of claim 62 wherein said one of said plurality of segments is the segment most temporally adjacent to said portion of said audio.

64.   The method of claim 62 wherein said one of said plurality of segments is the next temporally related segment.

65.   The method of claim 62 wherein said one of said plurality of segments is the previous temporally related segment.

66.   The method of claim 56 wherein a user selects a portion of said audio included within said plurality of segments, wherein in response thereto, said system presents said portion of said audio from the start thereof.

67.   The method of claim 56 wherein a user selects a portion of said audio not included within said plurality of segments, wherein in response thereto, said system presents one of said plurality of segments, and wherein said user   selects a portion of said audio included within said plurality of segments, wherein in response thereto, said system presents said portion of said audio within said plurality of segments.

68.   The method of claim 56 wherein a user selects a portion of said audio not included within said plurality of segments, wherein in response thereto, said system presents one of said plurality of segments, and wherein said user selects a portion of said audio included within said plurality of segments, wherein in response thereto, said system presents said portion of said audio within said plurality of segments starting from the beginning thereof.

69.   The method of claim 56 wherein a user selects a portion of said audio not included within said plurality of segments, wherein in response thereto, said system presents said selected portion not included within said plurality of segments, and wherein after presenting said selected portion not included within said plurality of segments presents said selected plurality of segments in temporal order without portions of said audio not included within said plurality of segments, and wherein said user selects a portion of said audio included within said plurality of

segments, wherein in response thereto, said system presents said portion of said audio within said plurality of segments.

70.   The method of claim 56 wherein said temporal information is hierarchical and is displayed in such a manner to retain a portion of its hierarchical structure.

71.   The method of claim 56 wherein said temporal information relates to overlapping time periods and said temporal information is displayed in such a manner to maintain the differentiation of said overlapping time periods.

72.   The method of claim 56 wherein said temporal information is displayed within a time line, wherein the temporal information is presented in a plurality of layers in a direction perpendicular to the length of said time line.

73.   The method of claim 56 wherein said temporal information is displayed within a time line, wherein textual information is included within said time line.

74.   The method of claim 56 wherein said temporal information is displayed within a time line, wherein additional textual information is displayed upon selecting a portion of said time line.

75.   The method of claim 56 wherein said temporal information is displayed together with a time line, wherein additional textual information is displayed together with selecting a portion of said time line.

76.   The method of claim 56 wherein said temporal information is displayed within a time line, wherein additional audio annotation is presented upon presenting a portion of said time line.

77.   The method of claim 29 wherein a user selectable skip function skips a set of frames to a modified location of said video in at least one of a forward temporal direction or a reverse temporal direction, and displays said video at said modified location.

78.   The method of claim 29 wherein a user selectable skip function skips to a later temporal segment or a previous temporal segment, and displays said video at said later temporal segment or said previous temporal segment, respectively.

79.   The method of claim 29 wherein a user selectable scan function skips a set of frames to a modified location of said video in at least one of a forward temporal direction or a reverse temporal direction, and displays said video at said modified location, and thereafter automatically skips another set of frames to another modified location of said video in at least one of said forward temporal direction or said reverse temporal direction, and displays said video at said another modified location.

80. The method of claim 79 wherein at least one of said forward temporal direction and said reverse temporal direction are consistent with said different segments.

81. The method of claim 80 wherein said display of said video is at the start of the respective one of said different segments.

82. The method of claim 80 wherein said display of said video is at a predetermined offset within the respective one of said different segments.

83. The method of claim 29 wherein said graphical user interface displays a respective image associated with at least a plurality of said different segments.

84. The method of claim 82 wherein said respective image associated with the currently presented said different segments is visually highlighted.

85. The method of claim 83 wherein during presentation of said video said visually highlighted respective images are said highlighted in a substantially regular interval while the sequence of said presentation of said video is at substantially irregular intervals.

86. The method of claim 56 wherein the presentation of said different segments may be modified by a plurality of different functions, and wherein the user may customize another function, not previously explicitly provided, by combining a plurality of said plurality of different functions into a single function.

EVIDENCE APPENDIX:

None.

RELATED PROCEEDINGS APPENDIX:

None.